

Draft Genome Sequence of *Paenibacillus* sp. Strain DMB5, Acclimatized and Enriched for Catabolizing Anthropogenic Compounds

Jenny Johnson,^a Binal Shah,^a Kunal Jain,^a Nidhi Parmar,^b Ankit Hinsu,^b Namrata Patel,^b Chaitanya G. Joshi,^b Datta Madamwar^a

Environmental Genomics and Proteomics Laboratory, UGC Centre of Advanced Study, PG Department of Biosciences, Sardar Patel University, Bakrol, Anand, Gujarat, India^a; Department of Animal Biotechnology, College of Veterinary Science and Animal Husbandry, Anand Agricultural University, Anand, Gujarat, India^b

J.J. and B.S. contributed equally to this work.

Here, we present the draft genome sequence of *Paenibacillus* sp. strain DMB5, isolated from polluted sediments of the Kharicut Canal, Vatva, India, having a genome size of 7.5 Mbp and 7,077 coding sequences. The genome of this dye-degrading bacterium provides valuable information on the microbe-mediated biodegradation of anthropogenic compounds.

Received 13 February 2016 Accepted 18 February 2016 Published 31 March 2016

Citation Johnson J, Shah B, Jain K, Parmar N, Hinsu A, Patel N, Joshi CG, Madamwar D. 2016. Draft genome sequence of *Paenibacillus* sp. strain DMB5, acclimatized and enriched for catabolizing anthropogenic compounds. *Genome Announc* 4(2):e00211-16. doi:10.1128/genomeA.00211-16.

Copyright © 2016 Johnson et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Chaitanya G. Joshi, cgjoshi@aau.in, or Datta Madamwar, datta_madamwar@yahoo.com.

Industrialization has become inevitable for socioeconomic development, but on the other hand, unacceptable industrial waste discharges impose a severe threat to civilization itself. The emerging level of pollution has accelerated the eradication of various natural resources of our planet Earth. Many contaminated vicinities are partially bioremediated by the natural attenuation processes, emphasizing the significance of the indigenous microbial population. Microbes that are catabolically diverse and flexible can endure the extremely varied conditions by inherently developing adaptive mechanisms. The exploitation of this microbial richness will enable us to explore the taxonomical mutualism and functional stability that prevail to allow microbial assemblages to coexist in such niches. Henceforth, we have distinguished a potent bacterium, *Paenibacillus* sp. strain DMB5, from the polluted sediments of Kharicut Canal, receiving toxic industrial effluents from the Vatva Industrial Estate, India. The bacterial isolate, in synergism with other bacteria, effectively degraded and metabolized azo dyes.

The genome of *Paenibacillus* sp. strain DMB5 was shotgun sequenced on an Ion Torrent platform using 318 Chip and 300-bp chemistry. The sequencing conferred 5,390,184 reads and 1.6 Gb of nucleotide bases, with an average read length of 312 bp. After quality filtering using PrinSeq-lite version 0.20.4 (1), the filtered reads were *de novo* assembled using the CLC Genomics Workbench software (CLC bio-Qiagen, Aarhus, Denmark), MIRA (2), and GS *de novo* assembler version 2.6. All three assemblies were then converged in CISA (3) to generate 174 contigs with a mean size of 42,857 bp and maximum length of 399,150 bp. The total genome size is 7,457,220 bp, with a G+C content of 51.2%.

The draft genome of *Paenibacillus* sp. DMB5 was annotated using the Rapid Annotations using Subsystems Technology (RAST) server (4) and the NCBI Prokaryotic Genome Annota-

tion Pipeline (PGAP) (http://www.ncbi.nlm.nih.gov/genome/annotation_prok). The annotations revealed 7,077 coding sequences, 89 tRNA genes, and 17 rRNA genes. Various subsystems of essential functions, like carbohydrate metabolism, amino acid and derivatives, protein metabolism, etc., were discovered using RAST. Remarkably, the presence of several significant genes/operators in the *Paenibacillus* sp. DMB5 genome were observed that included the genes pertaining to resistance to antibiotics and toxic compounds, such as genes encoding metal efflux pump proteins for heavy metals, like mercury, chromium, arsenic, copper, cadmium, and zinc. Many stress-related proteins were identified as being related to oxidative stress, heat shock, and detoxification. Within the oxidative stress category, NADH quinone oxidoreductases, monooxygenases, and dioxygenases were also annotated that are reported to catalyze the reduction of azo dyes (5-7). Moreover, a number of proteins involved in various aromatic compound degradation pathways were noticed, including pathways for quinate, salicylate, and gentisate degradation. The relevance of these catabolic genes in the draft genome validates the potential of *Paenibacillus* sp. DMB5 for the biodegradation of xenobiotic compounds. However, the pursuit to explore the underlying degradation mechanisms and biochemical pathways in strain DMB5 still remains.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. [LRAC00000000](https://www.ncbi.nlm.nih.gov/nuclink/LRAC00000000). The version described in this paper is version LRAC01000000.

ACKNOWLEDGMENTS

This work was supported by the Department of Biotechnology (DBT) grant BT/1/CEIB/09/V/05 from the Ministry of Science and Technology, New Delhi, India. J.J. acknowledges the Department of Science and Technology (DST), New Delhi, India, for INSPIRE fellowship.

FUNDING INFORMATION

This work, including the efforts of Datta Madamwar, was funded by Department of Biotechnology, Ministry of Science and Technology (DBT), New Delhi, India (BT/01/CEIB/09/V/05).

REFERENCES

- Schmieder R, Edwards R. 2011. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 27:863–864. <http://dx.doi.org/10.1093/bioinformatics/btr026>.
- Chevreur B, Wetter T, Suhai S. 1999. Genome sequence assembly using trace signals and additional sequence information, p 45–56. In *Computer science and biology. Proceedings of the German Conference on Bioinformatics, GCB '99*. GCB, Hannover, Germany.
- Lin SH, Liao YC. 2013. CISA: contig integrator for sequence assembly of bacterial genomes. *PLoS One* 8:e60843. <http://dx.doi.org/10.1371/journal.pone.0060843>.
- Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formisano K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: Rapid Annotations using Subsystems Technology. *BMC Genomics* 9:75. <http://dx.doi.org/10.1186/1471-2164-9-75>.
- Keck A, Klein J, Kudlich M, Stolz A, Knackmuss HJ, Mattes R. 1997. Reduction of azo dyes by redox mediators originating in the naphthalenesulfonic acid degradation pathway of *Sphingomonas* sp. strain BN6. *Appl Environ Microbiol* 63:3684–3690.
- Ryan A, Kaplan E, Nebel J, Polycarpou E, Crescente V, Lowe E, Preston GM, Sim E. 2014. Identification of NAD(P)H quinone oxidoreductase activity in azoreductases from *P. aeruginosa*: azoreductases and NAD(P)H quinone oxidoreductases belong to the same FMN-dependent superfamily of enzymes. *PLoS One* 9:e9855. <http://dx.doi.org/10.1371/journal.pone.0098551>.
- Shah B, Jain K, Patel N, Pandit R, Patel A, Joshi CG, Madamwar D. 2015. Draft genome sequence of *Paenibacillus* sp. isolated from Alang ship breaking yard, which harbors genes for xenobiotic degradation. *Genome Announc* 3(3):e00554-15. <http://dx.doi.org/10.1128/genomeA.00554-15>.